APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Apr. 2007, p. 2207–2217 Vol. 73, No. 7
0099-2240/07/$08.00+0 doi:10.1128/AEM.01599-06

Survey of Genomic Diversity among *Enterococcus faecalis* Strains by Microarray-Based Comparative Genomic Hybridization▽†

Ågot Aakra,* O. Ludvig Nyquist, Lars Snipen, Turid S. Reiersen, and Ingolf F. Nes

Department of Chemistry, Biotechnology and Food Science, The Norwegian University of Life Sciences, N-1432 Ås, Norway

Received 11 July 2006/Accepted 30 December 2006

We have compared nine *Enterococcus faecalis* strains with *E. faecalis* V583 by comparative genomic hybridization using microarrays (CGH). The strains used in this study (the "test" strains) originated from various environments. CGH is a powerful and promising tool for obtaining novel information on genome diversity in bacteria. By CGH, one obtains clues about which genes are present or divergent in the strains, compared to a reference strain (here, V583). The information obtained by CGH is important from both ecological and systematic points of view. CGH of *E. faecalis* showed considerable diversity in gene content: Compared to V583, the percentage of divergent genes in the test strains varied from 15% to 23%, and 154 genes were divergent in all strains. The main variation was found in regions corresponding to exogenously acquired or mobile DNA in V583. Antibiotic resistance genes, virulence factors, and integrated plasmid genes dominated among the divergent genes. The strains examined showed various contents of genes corresponding to the pTEF1, pTEF2, and pTEF3 genes in V583. The extensive transport and metabolic capabilities of V583 appeared similar in the test strains; CGH indicated that the ability to transport and metabolize various carbohydrates was similar in the test strains (verified by API 50 CH assays). The contents of genes related to stress tolerance appeared similar in V583 and the nine test strains, supporting the view of *E. faecalis* as an organism able to resist harsh conditions.

Enterococcus faecalis is a "two-faced" organism. Most strains of the species are harmless commensals. They may be beneficial to health (probiotic), they can be found naturally in raw material, and they can be used in starter cultures for fermented food. Other strains are feared opportunistic pathogens, causing serious illnesses in hospitals all over the world. Like many other strains of lactic acid bacteria, *E. faecalis* strains are intrinsically resistant to commonly used antibiotics and they acquire antibiotic resistance determinants rapidly. Because of the versatile nature of this bacterium as a commensal and as a pathogen and its impact on human health, it is of great interest to study the mechanisms, that direct the shift between the harmless and pathogenic conditions. Much information on this can be obtained by studies of the genetic diversity of the species. Until a few years ago, diversity of bacteria was often characterized by sequencing of rRNA genes, a limited number of other genes, and/or signature sequences. For determination of species affiliation, DNA reassociation experiments to determine DNA homology values are, despite many limitations, still the most important method. The use of genome-wide microarrays opens new possibilities regarding detailed information on the genomic diversity of bacteria, as well as the determination of species affiliations. For versatile organisms like *E. faecalis*, it is of interest to study which mechanisms lead to this versatility. By using microarrays to study genomic diversity in detail, one may obtain clues regarding the evolution of the strains within a species, e.g., how one strain becomes useful in foods while another strain becomes a feared pathogen.

In general, the knowledge of genomic diversity among bacteria is scarce, but the availability of complete genome sequences and the development of tools such as microarrays enable us to study whole genomes efficiently. It is believed that information on gene content and genome organization will give important clues about the driving forces in bacterial evolution. Some studies on the genetic and genomic diversity of enterococci have been published (5, 6, 14, 15, 18, 31), but only after the completion of a genome sequence is it possible to do genome-wide comparisons of strains within a species.

The complete genome sequence of *E. faecalis* V583 has been published (23). V583 is a vancomycin-resistant clinical isolate whose genome contains a large number of mobile genetic elements (MGEs). The acquisition of MGEs is believed to be an important mechanism by which the species *E. faecalis* has been able to generate genetic diversity and, thereby, highly variable phenotypes (23). Microarrays representing the V583 genome can be used to study transcriptional activities, and they can be applied in studies of diversity within the species on a genome-wide scale. By allowing hybridization of genomic DNA from an "unknown" strain to the microarray, one can obtain a measurement of the similarity and divergence between two closely related strains. This technique, comparative genomic hybridization (CGH), or genomotyping, is a powerful tool to estimate whole-genomic diversity and to determine the minimal core of functional genes in a bacterial species (19). CGH is also a relatively rapid method, although it is still expensive because of the costs associated with microarray experiments. CGH provides a promising predictive tool for correlating phenotypes and genotypes, and it permits genome scale analysis of the

* Corresponding author. Mailing address: Laboratory of Microbial Gene Technology, Department of Chemistry, Biotechnology and Food Science, The Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway. Phone: 47 6496 5894. Fax: 47 6494 1465. E-mail: agot.aakra@umb.no.

† Supplemental material for this article may be found at http://aem.asm.org/.

▽ Published ahead of print on 12 January 2007.

TABLE 1. Bacterial strains used in this study

E. faecalis strain[a]	Origin and/or characteristic(s)	Reference(s)
NCDO 581	Type strain	
NCDO 642	Type strain	
EF BRIDGE	Food	
JH2-SS	Laboratory; Str^r Spc^r pAD1	33
V24	Environment, bacteriocin producer	17
OG1RF	Laboratory; Rif^r Fus^r	21
INY3000	Tn916 mutant of OG1SSp	3, 34
138Vet	Veterinarian	
179Vet	Veterinarian, multiresistant	
V583	Hospital; Van^r	23, 27

[a] The strain designations shown are used as abbreviations for the strains throughout this paper.

molecular population genetics of a species. CGH has now been applied in studies of several bacterial species (see, e.g., references 2, 4, 7, 13, 28, and 32), and the CGH studies have given important insight into and indications about evolutionary events (4, 7, 13), genome composition (9), pathogenicity potential (16, 24), and horizontal gene transfer (19). In a recent study by Lepage et al. (18), a macroarray containing selected genes from *E. faecalis* V583 and MMH594 was used in a CGH approach related to food safety. Presumably, CGH with genome-wide microarrays of enterococci will provide crucial information about their evolution, survival mechanisms, and pathogenic potential.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The bacterial strains used in this study are listed in Table 1. All strains were grown overnight in flasks of brain heart infusion medium (Difco) at 37°C on a rotary shaker (300 rpm).

Microarrays, DNA extraction, labeling, and hybridization. The microarrays used have been described previously (1). DNA extraction was performed with ADVAMAX beads (EdgeBioSystems) as follows. Cell lysis was performed by resuspending cell pellets in a buffer containing 10% sucrose, 2 mg/ml lysozyme, 0.4 mg/ml RNase A, 25 mM Tris (pH 8.4), and 25 mM EDTA (pH 8) and incubation for 10 min at 37°C prior to addition of sodium dodecyl sulfate (SDS; final concentration, 1%) and NaCl (final concentration, 0.8 M) and incubation at 65°C for 5 min. Then, 100 μl ADVAMAX beads (~15% of the total volume) and $MgCl_2$ (final concentration, 0.09 M) was added, mixed by vortexing, and centrifuged for 5 min in a table-top centrifuge at maximum speed. The DNA was precipitated from the supernatant with isopropanol, dried, and resuspended in 1× Tris-EDTA (pH 8.0).

For the CGH experiments, 5 μg DNA from each strain was used. The DNA was digested with RsaI (Promega) and cleaned up with the QIAQuick PCR purification kit (QIAGEN) or the Bioprime DNA labeling kit (Invitrogen). Fluorescent labeling was performed by random priming with the Bioprime DNA labeling kit (Invitrogen). The concentrations of nonfluorescent nucleotides in the reaction mixtures were 0.12 mM each dATP, dCTP, and dGTP and 0.06 mM dTTP. To each labeling reaction mixture (50- or 25-μl total volume), fluorescent nucleotide analogs (cyanine 5-dUTP or cyanine 3-dUTP; Perkin-Elmer Life Sciences) were added to a final concentration of 0.06 mM. Reactions were stopped by addition of EDTA (pH 8.0) and cleaned up with Microcon-30 filters (Amicon/Millipore). Hybridizations to the microarrays were conducted as follows. Slides were prehybridized by incubation at 50°C for 30 min in a solution containing 1% bovine serum albumin (Calbiotech), 3.5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate), and 0.1% (wt/vol) SDS. Dried, fluorescently labeled samples were resuspended in a hybridization solution (40 μl) containing 5× SSC, 0.1% (wt/vol) SDS, 1.0% (wt/vol) bovine serum albumin, 50% (vol/vol) formamide, and 0.01% (wt/vol) single-stranded salmon sperm DNA. The resuspended probes were applied to the arrays underneath a lifterslip (Erie Scientific Co.) and incubated in the dark at 42°C for 6 h or overnight. Following hybridization, excess hybridization solution and unspecifically bound probe were washed away during four washing steps with gentle shaking, in the dark, as follows: 2 min in 2× SSC–0.1% SDS, 1 min in 1× SSC, 1 min in 0.2× SSC, and 30 s in 0.05× SSC. Immediately after washing, the arrays were dried by centrifugation at 600 rpm for 5 min in an Eppendorf 5810R table-top centrifuge (Eppendorf). Two replicate hybridizations (dye swap) were performed for all strains.

Hybridized arrays were scanned at wavelengths of 532 nm (cyanine-3) and 635 nm (cyanine-5) at a 10-μm resolution to obtain two TIFF images with a ScanArrayExpress Microarray Scanner (Packard Bioscience). Fluorescence intensities and spot morphologies were analyzed with QuantArray program ver. 3.0 (Packard BioScience) or GenePix Pro 6.0 (Molecular Devices), and spots were excluded on the basis of slide or morphology abnormalities.

Microarray data analysis. Data preprocessing followed a standard procedure with respect to background correction and filtering (25). Due to the highly skewed distribution of log ratios in microbial CGH experiments (many small and few large values), a standard normalization procedure, e.g., subtraction of a trend found by LOWESS smoothing log ratios over total signal intensity, is problematic. To avoid the assumption of a symmetric distribution of the invariant genes, and since all experiments were dye swapped, we implemented a dye swap normalization procedure. The dye effect on the log ratios for every gene was LOWESS smoothed over total signal intensity and then used for correction of all the original log ratios on both arrays in the dye swap pair. For the downstream analysis, only the average normalized log ratio for each gene was used.

In order to discriminate between present and divergent genes, we calculated ROTMIX scores for every gene as described by Snipen et al. (30). This score (p_i) can be seen as a posterior probability of divergence; i.e., if p_i is close to 1, gene i is likely to be divergent (absent) and if p_i is close to 0, it is likely to be present. If this score is ~0.5, the status of the gene is uncertain and it may be left unclassified. In order to avoid many false discoveries (false positives and false negatives), we only classified as divergent genes with p_i values larger than 0.9 and as present genes with p_i values smaller than 0.1.

API 50 CH for determination of fermentation patterns. The nine test strains and the reference strain, V583, were subjected to the API 50 CH assay (Bio-Mérieux, Marcy l'Étoile, France) according to the manufacturer's instructions. Cell suspensions were applied to API 50 CH strip wells, which were coated with petrolatum and incubated at 37°C. The results were read after 24 h and verified after 48 h. Fermentation of carbohydrates in the carbohydrate medium was indicated by a yellow color. Color reactions were scored against a chart provided by the manufacturer.

Microarray data accession number. The microarray data have been deposited in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress/) with the series accession number E-MEXP-931.

RESULTS AND DISCUSSION

The total number of genes from *E. faecalis* V583 represented on the microarrays was 3,160, which comprised both chromosomal (3,023 genes) and plasmid-carried genes (137 genes). Because of PCR failure during the construction of the V583 microarray, some of the genes (n = 177) were missing from our array analysis. It should be mentioned that we here use the term "gene" for sequences that actually are predicted genes or open reading frames (ORFs), according to Paulsen et al. (23) and the Comprehensive Microbial Resource (http://cmr.tigr.org/tigr-scripts/CMR/GenomePage.cgi?org_search=&org=gef), of which the majority have not been experimentally verified to be functional genes.

A divergent gene may be a gene which is absent in the test strain or a gene whose sequence has evolved so much that detectable hybridization cannot be obtained. Genes that obtained p_i scores between 0.1 and 0.9 were left unclassified in this study. A histogram (Fig. 1) showing a typical distribution of p_i (ROTMIX) scores justifies our choice of cutoff values (0.1 and 0.9). As shown in Fig. 1, this is actually a large majority of the genes in the array. The histogram shows that the choice of cutoff values is not critical (cutoffs of 0.25 and 0.75 would not give very different results). At around 0.5, there are a number of "difficult" genes, for which classification will be uncertain. By not classifying these genes, we avoid a number of false

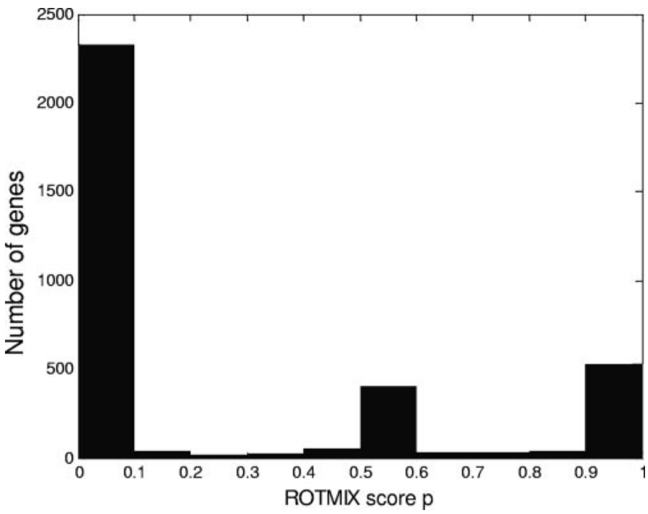

FIG. 1. Histogram of typical ROTMIX (p_i) score distribution. This is for the hybridizations with strain NCDO 642.

statements. A relatively low number of genes fell into the unclassified category. The highest number of such genes (chromosomal) was found in strain 138Vet (175 genes), while in strain NCDO581 there were only 54 unclassified genes. In this work, our primary focus was a comparison of the presence or divergence of V583-like chromosomal genes in the other nine strains. But on the basis of this CGH analysis, it is not possible to decide whether a gene is present on a chromosome or on a plasmid in a test strain.

The fundamental assumptions behind the use of the ROTMIX score for data analysis, as described by Snipen et al. (30), is that the microarray data are distributed in a V-shaped pattern when displayed in a microarray plot. This was indeed the case for the present data (results not shown). Hence, for the data obtained here, we expected this procedure to work very well. The unclassified genes in a data set are typically found at low microarray signal intensities (not shown).

We found that 154 chromosomal genes were divergent in all nine of the strains examined while 1,125 genes were classified as divergent in one or more of the nine strains studied. Out of the 154 genes that were divergent in all of the strains, more than 100 are classified as genes encoding hypothetical proteins. For 45 of these genes, a function has been predicted, and these genes are listed in Table 2. Compared to the V583 chromosome, the percentage of divergent genes in the test strains varied from 15% in NCDO 581 to 23% in INY3000.

The results of our genome-wide CGH supported the results of the macroarray CGH by Lepage et al. (18). Similar patterns of gene content variability were found in our work. The suggestions of Lepage et al. (18) on the distribution of selected genes between food and clinical isolates were also supported. Genes shared by all strains may be part of the core of functional genes defining the species *E. faecalis*. With a few exceptions, the 124 genes found to be conserved in the work of Lepage et al. (18) were also conserved (present in all strains) in our work. The gene EF3041, coding for a pheromone-binding protein, was divergent in all of the strains analyzed by us, while the gene was present in all of the strains analyzed by Lepage et al. (18).

MGEs: gene content in test strains. The main variation in gene content revealed by the CGH analysis relates to the mobilome (i.e., the pool of MGEs in a prokaryotic genome [8]) of V583, as illustrated in Fig. 2 and 3. The pattern of the presence or divergence of genes in the test strains is shown in Fig. 2. Comparison of the nonmobile gene pool and the DNA that has probably not been horizontally transferred showed that the similarities between the strains examined are high. In Fig. 2, the main regions of variability are highlighted. These regions correspond to putative MGEs that have been predicted in V583 (summarized by Lepage et al. in reference 18). The phage02 region, corresponding to EF1276 to EF1293, is not highlighted in Fig. 2, since nearly all phage02 genes were found to be present in the test strains; only the endolysin gene (EF1293) was divergent in one strain (OG1RF), and a few genes were unclassified. The phage02 region therefore appears

TABLE 2. A selection of genes that were divergent in all of the strains examined in this study[a]

ORF	Gene	Function
EF0137		Nucleotidyltransferase domain protein
EF0150		Membrane protein, putative
EF0166		Site-specific recombinase, phage integrase family
EF0304		Lipoprotein, putative
EF0306		Transcriptional regulator, Cro/CI family
EF0312		Aspartate 1 decarboxylase domain protein
EF0317		Transcriptional regulator, Cro/CI family
EF0510	*ssb-3*	Single-strand binding protein
EF0518		Cell wall surface anchor family protein
EF1442		DNA topoisomerase domain protein
EF2096		Tail protein
EF2115		Transcriptional regulator, ArpU family
EF2129		DNA replication protein, putative
EF2142		Transcriptional regulator, Cro/CI family
EF2144		Lipoprotein, putative
EF2145		Site-specific recombinase, phage integrase family
EF2255		Site-specific recombinase, phage integrase family
EF2259		Phosphosugar-binding transcriptional regulator, putative
EF2265		Carbohydrate kinase, pfkB family
EF2266		2-Dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase, putative
EF2269		PTS, IID component
EF2273		Transcriptional regulator, GntR family
EF2278		Lipoprotein, NLP/P60 family
EF2279		Membrane protein, putative
EF2286		ParB-like nuclease domain protein
EF2290		RNA polymerase sigma 70 factor, ECF subfamily
EF2291		Transcriptional regulator, Cro/CI family
EF2296	*vanW*	Vancomycin B-type resistance protein VanW
EF2300		Streptomycin resistance protein, putative
EF2323		Modification methylase, putative, truncation
EF2326		Group II intron reverse transcriptase maturase
EF2340		C-5 cytosine-specific DNA methylase
EF2350		Transcriptional regulator, Cro/CI family
EF2803		Holin
EF2955		Site-specific recombinase, phage integrase family
EF3034		Transcriptional regulator, GntR family
EF3037	*pepA*	Glutamyl-aminopeptidase
EF3041		Pheromone binding protein
EF3043		PTS, IIC component
EF3044	*nagA2*	*N*-Acetylglucosamine-6-phosphate deacetylase
EF3050	*tag-2*	DNA-3-methyladenine glycosylase I
EF3099		Transporter accessory protein, putative
EF3103		Membrane protein, putative
EF3104		ABC transporter, ATP-binding protein
EF3226		Rep protein

[a] Genes encoding hypothetical proteins which were deemed divergent in all strains are not listed.

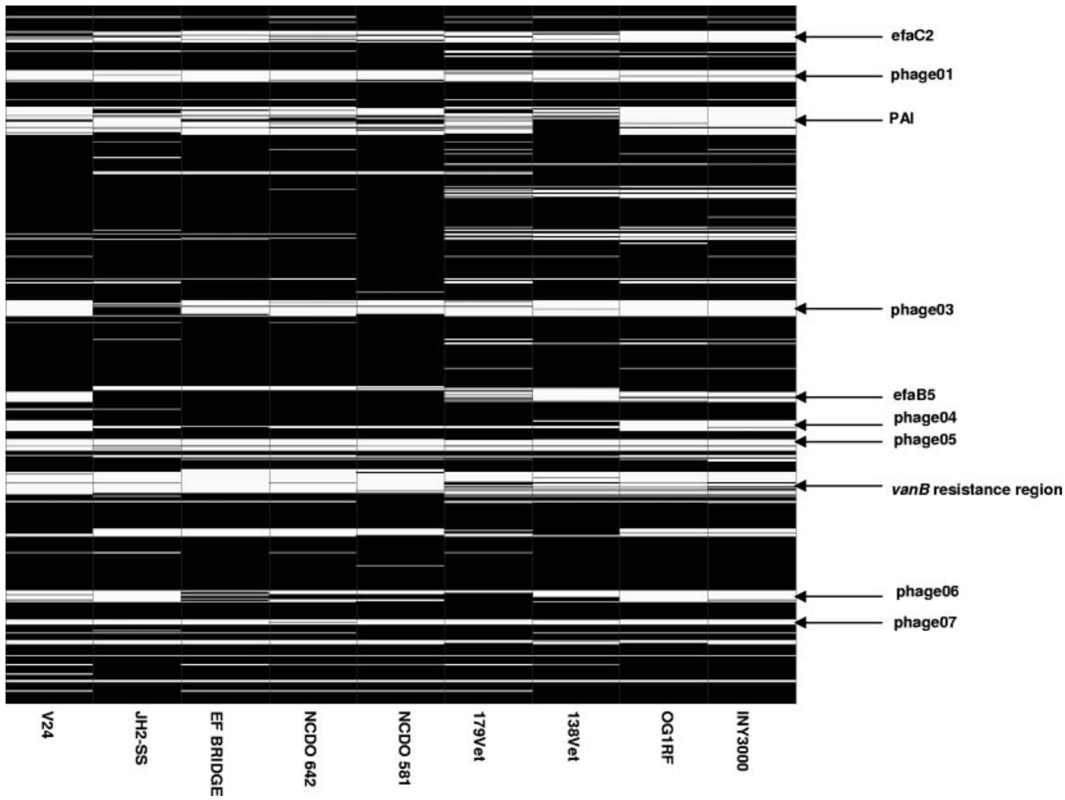

FIG. 2. Genome comparison of *E. faecalis* strains with the sequenced V583 strain. Each row corresponds to a specific spot (gene) on the microarray. The genes are arranged according to the numbering of V583 genes, with EF0001 at the top and EF3333 at the bottom. Each column corresponds to a test strain (Table 1). The status of each gene is indicated as follows: black, present; white, divergent; grey, unclassified. Positions of putative MGEs predicted in the V583 chromosome are indicated by arrows.

widespread and may be considered an ancient MGE of *E. faecalis*.

A summary of the content of genes corresponding to the MGEs of V583 (mentioned by Lepage et al. in reference 18) is shown in Fig. 3. The Tn*1549*-like transposon carries the vancomycin resistance-encoding genes (EF2283 to EF2334). Among the genes that were divergent in all of the test strains was EF2296, encoding the vancomycin resistance type B protein VanW, whose function in vancomycin resistance is not known. The other genes related to vancomycin resistance (EF2293 to EF2299) were divergent or unclassified in at least six test strains (Fig. 3). Hence, in the majority of the test strains we cannot claim that *vanB*-type vancomycin resistance is present. On the basis of the patterns of variability of the MGE genes in the test strains, one may hypothesize about the dissemination of the MGEs in the species and whether they are ancient or recent acquisitions. According to our CGH results, efaB5 (EF1857 to EF1895), phage04 (EF1988 to EF2043), and phage06 (EF2798 to EF2855) are ancient acquisitions. These MGEs display a higher level of gene presence in the test strains than the other MGEs do (Fig. 3). efaB5 contains pantothenate biosynthesis genes EF1859 and EF1860, which are discussed below. phage04 and phage06 both contain endolysin-encoding genes. Regarding the pathogenicity island (PAI), the highest level of gene divergence was found in strains V24, OG1RF, and INY3000, while 138Vet, which is a multiresistant veterinary isolate, had a high level of present PAI genes (Fig. 3).

Nearly 40 IS elements have been identified in V583 (23), and there are two clusters of IS elements. One of these clusters is associated with the PAI (see below) and contains integrated plasmid genes, while the other IS cluster flanks a region with an unusual nucleotide composition (EF1858 to EF1860, encoding *panDCB*; see below). The V583 PAI was described previously (29). It comprises the region from EF0479 to EF0628 in the V583 genome and contains several virulence factors. The PAI is a "mosaic" of several IS elements and integrated plasmid genes. Thus, as expected, the content of PAI genes in the strains examined varies.

Genes conserved among low-GC gram-positive bacteria. A comparison of the V583 genome sequence with 10 other sequenced low-GC gram-positive bacteria (23) revealed more than 500 conserved ORFs, constituting one-sixth of the genes in the V583 genome (see Table S1 in the supplemental material). These genes are mainly found outside the regions containing mobile or exogenously acquired DNA. A few of the conserved genes are located within the PAI of V583 or within other MGEs, which indicates that, in an evolutionary context, parts of the V583 mobilome are ancient (see Table S1 in the supplemental material). However, it is surprising that we found parts of an MGE conserved across these organisms. This probably relates to the fact that a majority of the 10 genomes with which V583 was compared are well-known pathogens (e.g., *Staphylococcus aureus*, *Listeria monocytogenes*, and *Streptococcus pneumoniae*). By the CGH approach, we found, as ex-

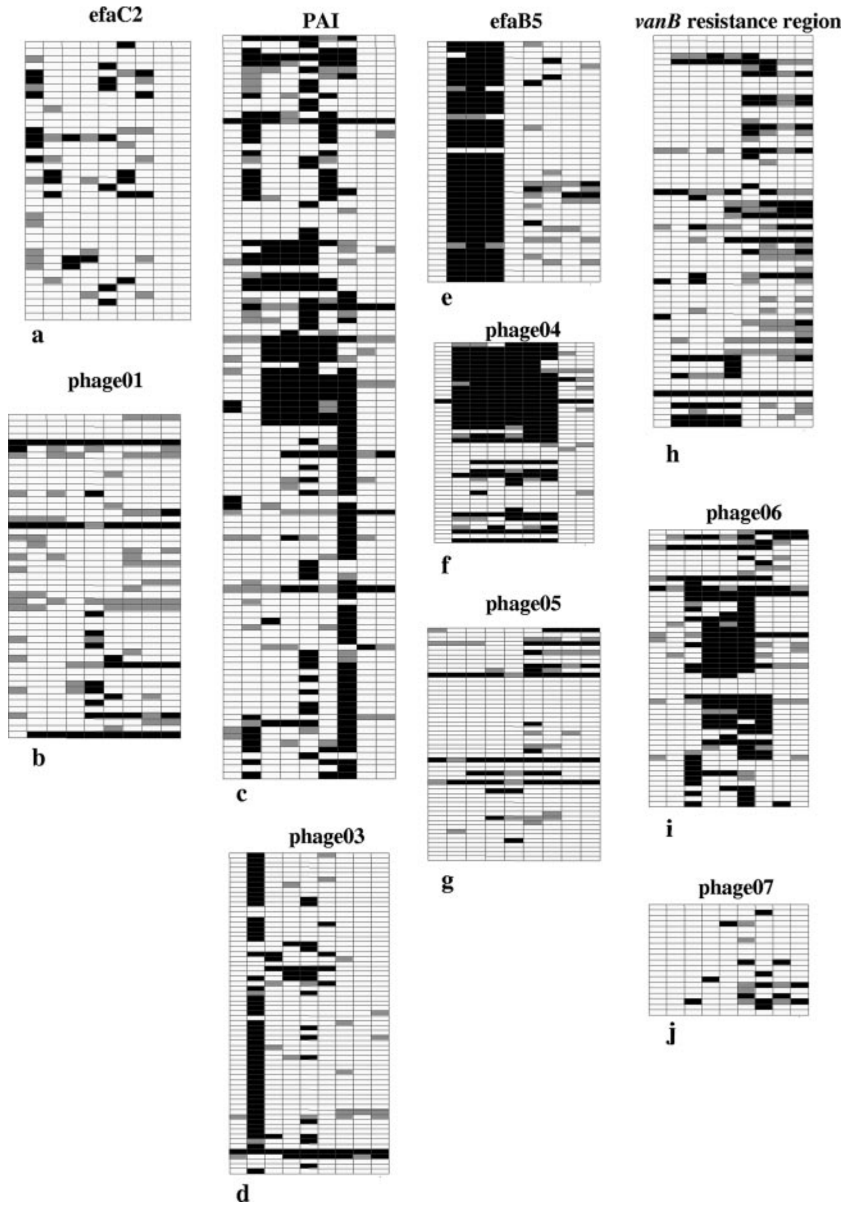

FIG. 3. Patterns of presence or divergence of MGE genes in test strains compared to V583. MGEs efaC1 and phage02 (see reference 18) are not shown. Panels: a, efaC2; b, phage01; c, PAI; d, phage03; e, efaB5; f, phage04; g, phage05; h, *vanB* resistance region; i, phage06; j, phage07. Each column in each panel corresponds to a test strain (Table 1), and in all panels the strains are listed equally, from the left, as follows: V24, JH2-SS, EF-BRIDGE, NCDO 642, NCDO 581, 179Vet, 138Vet,OG1RF, and INY3000. The status of each gene is indicated as follows: black, present; white, divergent; grey, unclassified. The genes are arranged according to the numbering of V583 genes as follows: efaC2, EF0127 to EF0166; phage01, EF0303 to EF0355; PAI, EF0479 to EF0628; phage03, EF1417 to EF1489; efaB5, EF1847 to EF1895; phage04, EF1988 to EF2043; phage05, EF2084 to EF2145; *vanB* resistance region, EF2277 to EF2346; phage06, EF2798 to EF2855; phage07, EF2936 to EF2955.

pected, that a large majority of the conserved genes were present in the test strains (see Table S1 in the supplemental material). Only seven of these conserved genes were divergent in all of the strains (see Table S1 in the supplemental material). These genes were EF0584 (coding for an ABC transporter, ATP-binding protein; unclassified in JH2SS), EF2257 (coding for a putative phosphotransferase system [PTS] IIC component; unclassified in NCDO581 and138Vet), EF2263 (coding for a putative gluconate 5-dehydrogenase; unclassified in 138Vet), EF2299 (coding for DNA-binding response regulator VanRB; unclassified in 138Vet), EF3037 (coding for a glutamyl aminopeptidase), EF3044 (coding for an *N*-acetylglucosamine-6-phosphate deacetylase), EFA0067 (coding for PTS IIABC components). Our CGH analysis confirms the presence of the majority of those genes considered conserved among low-GC gram-positive bacteria also in the *E. faecalis* strains examined. Fastidious organisms like *E. faecalis* are dependent on obtaining many nutrients from their environment, and accordingly, they possess a large amount of genes involved in the transport of nutrients into the cells. Genes coding for, e.g., ABC transporters and PTS components thus constitute a high percentage of the genes conserved among the low-GC gram-positive bacteria. The CGH analysis supported the notion that most of these genes were present in the species *E. faecalis* as a whole. Included among the conserved genes, we found, as expected, housekeeping genes like those encoding ribosomal proteins, translation factors, DNA polymerase, gyrase, ligase, and topoisomerase and acetyl coenzyme A carboxylase.

The importance of two-component signal transduction systems (TCSs) for the ability of *E. faecalis* to respond to environmental stimuli has been stressed by, e.g., Hancock and Perego (11, 12). Several genes encoding response regulators and histidine kinases are conserved among the low-GC gram-positive bacteria, and all of these genes (except EF2299, *vanRB*, and EF0571, within the PAI) were found to be present in the nine strains examined by CGH as well (see Tables S1 and S2 in the supplemental material). A further survey of the content of TCSs in the test strains follows below.

Although the comparison of V583 with these low-GC gram-positive bacteria showed a high degree of conservation between them and V583, the level of synteny was low (23). CGH microarrays do not reveal the locations of various genes within the test strains; thus, we cannot estimate the level of synteny among the strains we have studied here.

Transport and metabolic capabilities as revealed by CGH and API 50 CH. According to the gene content in the V583 genome, its transport and metabolic capabilities are extensive. In addition to 35 probable PTSs for sugar transport, the genome contains several ABC-type and other sugar transport systems, and genes encoding pathways for the utilization of 15 sugars have been predicted in the V583 genome (23). In general, the contents of PTSs and ABC-type sugar transporters appeared similar in V583 and the nine strains with which it was compared by CGH (summarized in Table 3). We suppose that the presence or divergence of the genes related to the various pathways indicates the ability of a strain to utilize a certain sugar for growth. To verify the CGH results for these genes, we used the API 50 CH assay to test the fermentation patterns of each strain.

The results suggested that all strains (including V583) have the ability to utilize cellobiose, fructose, lactose, galactose, glucose, glycerol, maltose, mannitol, mannose, *N*-acetylglucosamine, ribose, sucrose, and trehalose. Arabinose may be used by strains NCDO 581, V24, and EF BRIDGE. The API 50 CH test for rhamnose was positive for strains NCDO 581, NCDO 642, and 179Vet, while the genes needed for utilization of rhamnose was present in all of the strains tested. Thus, it appears that rhamnose fermentation is not a common property in *E. faecalis*, possibly because of a lack of transport capabilities for this sugar. The ability to use xylose appeared limited among the strains studied here; the *xylA* gene was present in 138Vet only. The API CH 50 assay supported this finding; none of the strains showed the ability to ferment xylose. In Table 3, the API 50 CH test results for the carbohydrates mentioned above, along with CGH results for selected genes related to transport and utilization of the sugars, are shown.

Virulence-related genes. These days, much attention is drawn to the pathogenicity of *E. faecalis*. The mechanisms by which the pathogenic strains of *E. faecalis* invade their hosts are not well known. The genome sequence of V583 revealed several groups of genes that may be related to its pathogenicity, e.g., genes that may enable adhesion to host cells and translocation across intestinal epithelium and surface-exposed putative virulence genes (23). A relatively high proportion of these genes was found to be present in all of the nine strains examined by CGH (see Table S3 in the supplemental material). Only 13 of these genes are found among the genes conserved among low-GC gram-positive bacteria. One may therefore speculate whether some of these surface-exposed, virulence-related genes are specific for the *E. faecalis* species. It might be that the products of such *E. faecalis* genes are important for the survival and growth of all *E. faecalis* strains. In the absence of certain virulence genes (or, rather, when they are not expressed), the strains become harmless, while the presence (expression) of certain other genes in the group contributes to the pathogenicity of other strains. Several suspected virulence genes are found within MGEs. Such suspected genes are EF0149 (in MGE efaC2), EF0485 (in PAI), EFA0047 (all coding for aggregation substance), the cytolysin operon EF0525 to EF0528 (in PAI), EF0700 (coding for hemolysin), EF0799 (coding for autolysin), and EF1992 (coding for endolysin in phage04). Several of these genes were present in more than one of the test strains (see Table S3 in the supplemental material). CGH results of regions comprising some of these genes are highlighted in Fig. 4. The hemolysin- and autolysin-encoding genes were present in all of the test strains (Fig. 4; see Table S3 in the supplemental material), while the endolysin gene (EF1992) was unclassified in OG1RF and divergent in V24 and INY3000 only. The content of the cytolysin operon and the aggregation substance genes was more variable (Fig. 4; see Table S3 in the supplemental material). Similar to what was found in our study, the presence of several virulence factors in harmless strains was observed also by Lepage et al. (18). This indicates that *E. faecalis* virulence is not a result of single virulence factors.

Stress resistance genes. *E. faecalis* is known as an organism able to tolerate highly variable growth conditions, and it tolerates high stress levels (see, e.g., reference 1 and references therein). In the V583 genome, these abilities are reflected in a number of genes (ca. 50) assumed to be involved in various types of stress resistance (e.g., resistance to oxidative stress, osmotic stress) (23). Of these predicted stress resistance genes, only 13 are conserved among the low-GC gram-positive bacteria mentioned by Paulsen et al. (23). Among the nine strains examined by CGH, only five genes appeared divergent in some of the strains (see Table S4 in the supplemental material). Thus, the predicted broad stress resistance capabilities appear to be widespread in the species *E. faecalis* and support the view of *E. faecalis* as a bacterium able to tolerate diverse and harsh conditions.

Two-component signal transduction systems. As mentioned above, TCSs in V583 have been analyzed and reviewed by Hancock and Perego (11, 12). The presence or divergence of the various TCSs in the strains we studied by CGH is shown in Table S2 in the supplemental material. Overall, the content of TCSs in the strains examined appeared similar to that of V583. The TCS located within the PAI (EF0570 and EF0571) was divergent in nearly all of the strains (see Table S2 in the supplemental material). Likewise, the TCS located within the mobile element containing the vancomycin resistance genes

TABLE 3. CGH results for selected genes related to transport and utilization of sugars and API 50 CH test results for selected carbohydrates

Carbohydrate(s)	ORF (gene)	Result for strain:																		
		V583,[a] API[b]	NCDO 581		NCDO 642		EF BRIDGE		JH2-SS		V24		OG1RF		INY3000		138Vet		179Vet	
			CGH[c]	API	CGH	API	CGH	API	CGH	API	CGH	API	CGH	API	CGH	API	CGH	API	CGH	API
Arabinose	EF1131 (*araD*)	Neg	+	Pos	ND	Neg	ND	Pos	+	Neg	+	Pos	−	Neg	−	Neg	−	Neg	ND	Neg
Cellobiose	EF1159	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF1160		+		+		+		+		+		+		+		+		+	
Fructose	EF0694	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF0717		+		+		+		+		+		+		+		+		+	
	EF1167 (*fba*)		+		+		+		+		+		+		+		+		+	
	EF1504		+		+		+		+		+		+		+		+		+	
Lactose and galactose	EF1834 (*lacB*)	Pos	−	Pos	−	Pos	−	Pos	−	Pos	+	Pos	+	Pos	+	Pos	−	Pos	−	Pos
	EF1835 (*lacA*)		−		−		−		−		+		+		+		ND		−	
Glucose	EF1416 (*pgi*)	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	−	Pos	−	Pos	−	Pos	ND	Pos
	EF1129		+		+		ND		−		ND		−		−		−		ND	
Glycerol	EF1928	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF1929 (*glpK*)		+		+		+		+		+		+		+		+		+	
Maltose	EF0954-EF0958	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
Mannitol	EF0411-EF0413	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
Mannose	EF0020	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF0021		+		+		ND		−		ND		−		−		−		−	
	EF0022		+		+		+		+		+		+		+		+		+	
	EF2589 (*manA*)		+		+		+		+		+		+		+		+		+	
N-Acetylglucosamine	EF0451	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF0466 (*nagB*)		+		+		+		+		+		+		+		+		+	
	EF3044 (*nagA2*)		−		−		−		−		−		−		−		−		−	
Rhamnose	EF0433	Neg	+	Pos	+	Pos	+	Neg	+	Neg	+	Neg	+	Neg	+	Neg	+	Neg	+	Pos
	EF0434 (*rhaA*)		+		+		+		+		+		+		+		+		+	
	EF0435 (*rhaD*)		+		+		+		+		+		+		+		+		+	
	EF1243		+		+		+		+		+		+		+		+		+	
Ribose	EF0197 (*rpiA*)	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	−	Pos	−	Pos	+	Pos	+	Pos
	EF1922		+		+		+		+		+		+		+		+		+	
	EF2959		+		+		+		+		+		+		+		+		+	
	EF2960 (*rbsD*)		+		+		+		+		+		+		+		+		+	
	EF2961 (*rbsK*)		+		+		+		+		+		+		+		+		+	
Sucrose	EF1603 (*scrB1*)	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF1746 (*galU*)		+		+		+		+		+		+		+		+		+	
	EFA0069 (*scrB2*)		−		−		−		−		−		−		−		−		−	
Trehalose	EF0955 (*galM*)	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos	+	Pos
	EF0956 (*pgmB*)		+		+		+		+		+		+		+		+		+	
	EF1068 (*galM*)		+		+		+		+		+		+		+		+		+	
	EF2788 (*glcK*)		+		+		+		+		+		+		+		+		+	
Xylose	EF0556 (*xylA*)	Neg	−	Neg	−	Neg	−	Neg	−	Neg	−	Neg	−	Neg	−	Neg	+	Neg	−	Neg

[a] Only API 50 CH results are shown for V583 since all of the genes, by definition, were present in this strain.
[b] Fermentation of carbohydrates in the carbohydrate medium was indicated by a yellow color. Pos indicates fermentation, and Neg indicates no reaction, no color change.
[c] +, gene present; −, gene divergent; ND, gene unclassified.

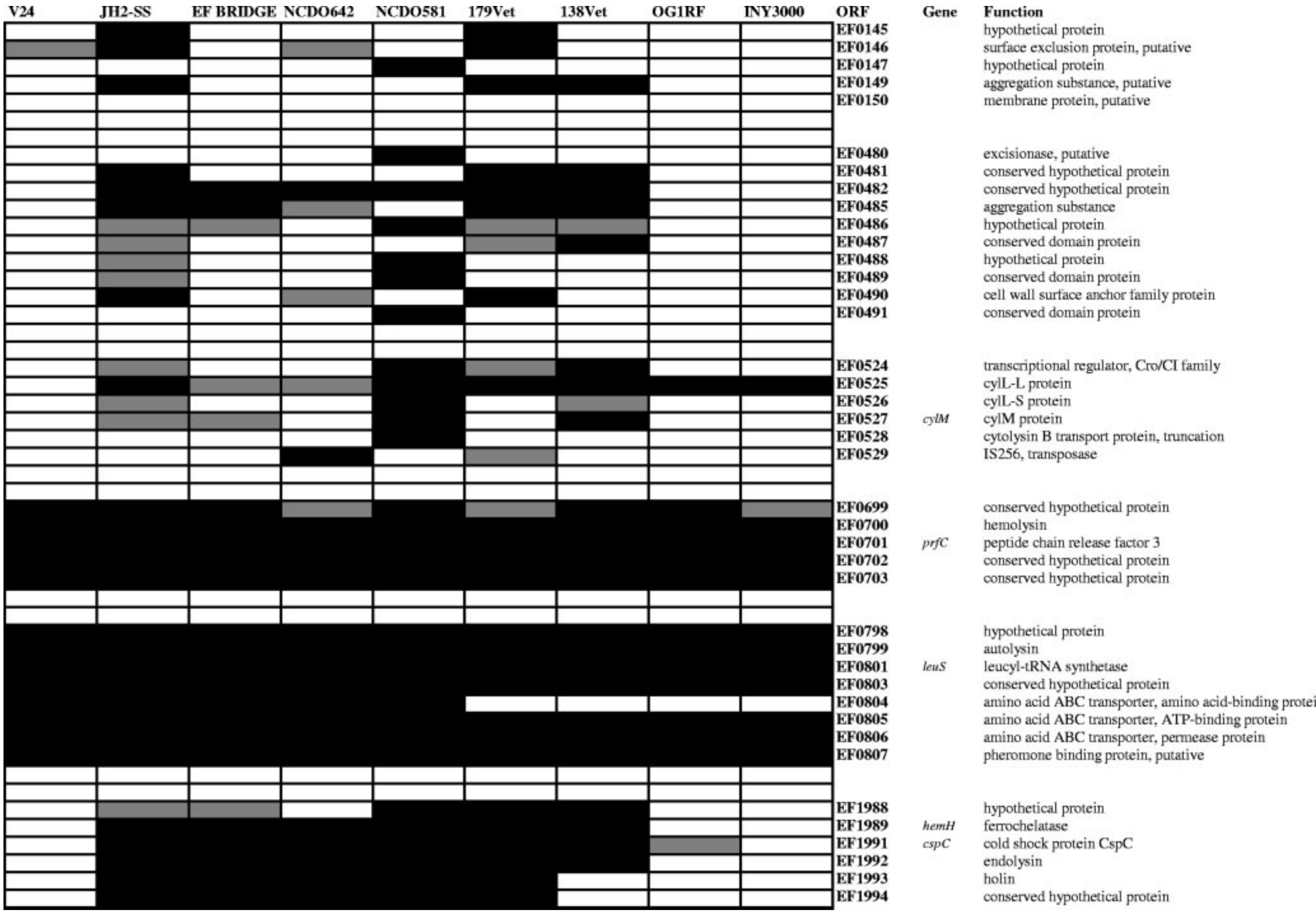

FIG. 4. CGH results for selected regions of the V583 genome containing virulence-related genes (see text). The status of each gene is indicated as follows: black, present; white, divergent; grey, unclassified.

was divergent (EF2298 and EF2299; see Table S2 in the supplemental material). The TCS encoded by EF1335 and EF1336 appeared divergent in V24, JH2SS, 179Vet, OG1RF, and INY3000, while the TCS encoded by EF1863 and EF1864 was divergent in V24, 179Vet, 138Vet, OG1RF, and INY3000 and the gene for the sensor histidine kinase (EF1863) was unclassified in JH2SS and EF BRIDGE. The EF1863-EF1864 region encodes the *vncSR* TCS (mentioned above) and is located within a mobile element. The EF1869-to-EF1863 region is homologous to a region encoding the VncRS two-component system and Vex secretion proteins in *S. pneumoniae*, and it has been associated with vancomycin tolerance in the latter bacterium (20, 26). The EF1869-to-EF1863 region is flanked by IS*256* (the efaB5 MGE), so one may assume that it is horizontally transferred (23). It was expected to be divergent in one or several of the strains examined. EF1335 and EF1336 have been predicted to play a role in the uptake of macrolide antibiotics, like erythromycin, in V583. It is interesting that this TCS appeared divergent in the multiresistant 179Vet strain. The TCS was divergent in OG1RF, INY3000, V24, and JH2SS, as well (Table 2). However, the EF1335 gene was found to be conserved (present) in all of the strains analyzed by Lepage et al. (18), and the distribution of this gene should be further examined.

Regarding transcriptional regulators other than TCSs, most genes coding for transcription factors were present in the strains studied. (The genes for sigma factors, EF0049, EF0782, EF1522, and EF3180, found in V583, were present in all of the strains.) An exception is genes coding for transcriptional regulators in the Cro/CI family. The V583 genome contains 36 such regulator genes, and of these, 24 were divergent in at least six strains. Three of these are located in the PAI or the vancomycin resistance MGE in V583, and two are located in the pTEF2 plasmid. However, genes coding for transcription factors in the Cro/CI family are often phage related, and the divergence of a majority of these genes in the examined strains may reflect the high content of phage DNA within the V583 genome.

Genes for pantothenate biosynthesis. Pantothenate is a precursor of coenzyme A, and thus, this compound is essential in all organisms. Few organisms, though, are able to synthesize pantothenate. It is assumed that V583 has acquired the pantothenate biosynthesis genes from other bacteria (streptococci, lactococci, clostridia) (23). Nine genes (EF0514, EF0517, EF0521, EF0522, EF1655, EF1858 to EF1860, and EF2445) in V583 have been predicted to be involved in pantothenate biosynthesis. Among these are three paralogs of the 2-dehydropantoate 2-reductase (23). Two of the paralogs (EF1655,

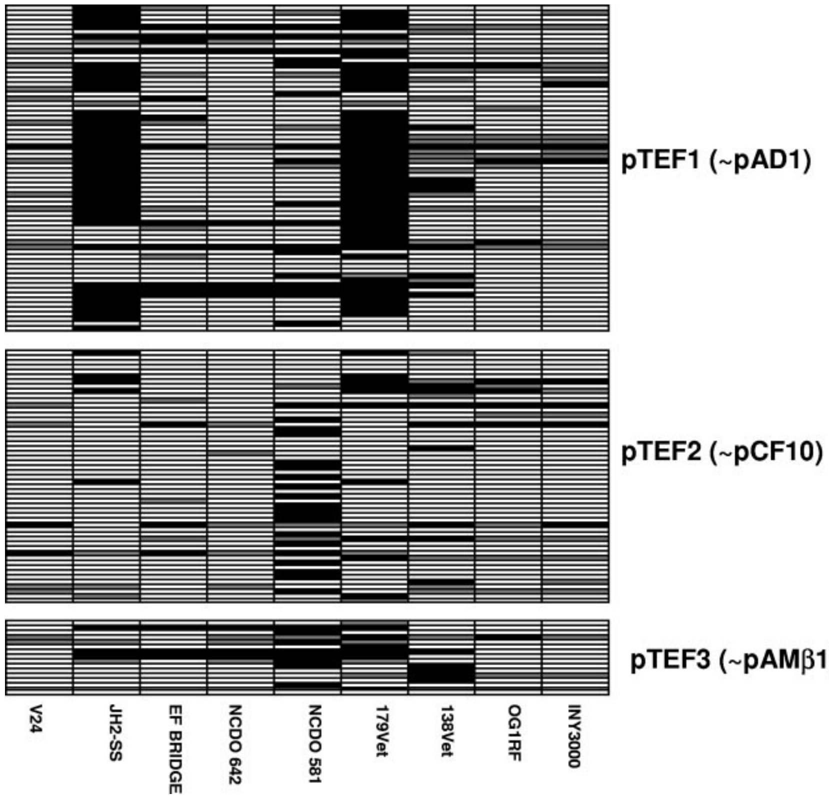

FIG. 5. Comparison of *E. faecalis* strains with respect to the content of genes corresponding to the genes in plasmids pTEF1, pTEF2, and pTEF3 of V583. The genes are arranged according to the numbering of V583 genes. Each column corresponds to a test strain. The status of each gene is indicated as follows: black, present; white, divergent; grey, unclassified.

EF2445) were present in all of the strains examined. The nine genes associated with pantothenate biosynthesis were all found to be present in JH2SS, EF BRIDGE, NCDO 642, and NCDO 581. EF0514, EF0517, EF0521, and EF0522 probably belong to a transposon (part of the PAI), and the genes appeared divergent in four strains (V24, 138Vet, OG1RF, and INY3000), while in 179Vet three of these genes were present (EF0517; divergent). V24, 138Vet, OG1RF, and 179Vet lack genes EF1858 to EF1860 (within the efaB5 MGE), which are presumed to have been acquired from clostridia by V583 (23). The variation in the content of the pantothenate biosynthesis genes indicates that the acquisition in V583 of some of the genes came about by relatively recent evolutionary events.

The ferrochelatase gene. EF1989 is a ferrochelatase gene (*hemH*) of probable phage origin (phage04) which may enable V583 to use coproporphyrinogen III for heme biosynthesis (23). *hemH* is divergent in V24, OG1RF, and INY3000, while the other strains may have the ability to use coproporphyrinogen III for heme biosynthesis in common with V583. Considering the fact that V24 is an environmental strain and OG1RF and INY3000 are laboratory strains, this is a reasonable finding.

Content of plasmid genes in the *E. faecalis* strains examined by CGH. Three regions of plasmid genes integrated into the V583 chromosome were identified by Paulsen et al. (23), of which one region is located within the PAI. Two of these plasmid regions contain genes encoding the virulence-associated aggregation substance. Genes encoding lipoproteins with signal peptides resembling those of pheromone precursors are also found within the regions of integrated plasmid genes (23). In general, most of the integrated plasmid genes were divergent in the strains examined. The main exceptions were as follows. The plasmid genes within the PAI (EF0485 to EF0506) appeared present in JH2-SS and 179Vet. The efaC1 region (EF2512 to EF2546; Lepage et al. [18]) was present in 138Vet and partly present in V24. Thus, the integrated plasmid genes add to the plasticity of the enterococcal genomes, and the presence or divergence of such regions may reflect the importance of these genes in the original habitat of each strain.

The genome of V583 comprises, in addition to the chromosome, three plasmids, of which two are pheromone responsive. These plasmids have been designated pTEF1, which is similar to the pAD1 plasmid; pTEF2, which is similar to pCF10; and pTEF3, which is like pAMβ. The strains examined here by CGH showed various contents of these plasmid genes (Fig. 5).

As expected, the majority of the pTEF1 genes (48 of the 68 genes examined) appeared present in JH2SS (Fig. 5), which harbors the pAD1 plasmid. Likewise, 179Vet probably harbored a pTEF1-like plasmid (Fig. 5); 52 of the 68 genes were classified as present. In the other strains, the contents of pTEF1 genes varied (Fig. 5) from 1 (EFA0038; V24) to 17 (NCDO 581). The unique parts of pTEF1 comprise a transposon with an aminoglycoside resistance gene and an IS-flanked element with other drug resistance genes. These pTEF1-specific genes (e.g., EFA0007; *ermB*, the gene for the rRNA adenine dimethylase family protein; EFA0010; and

EFA0061, the gene for a multidrug resistance protein) appeared divergent in most of the test strains; EFA0007 was divergent in all of the test strains; EFA0010 was present in EF BRIDGE and unclassified in JH2SS, NCDO642, and 170Vet; and EFA0061 was present in 179Vet only.

Regarding the genes of the pCF10-like plasmid pTEF2 (Fig. 5), the highest content of present genes was found in strain NCDO 581 (20 of 53 genes present). For the other strains, the number of present pTEF2 genes varied from 0 (NCDO 642) to 10 (179Vet). The novel pheromone inhibitor on pTEF2 (EFB0007) appeared divergent in all of the test strains (Fig. 5).

pTEF3 is the smallest of the plasmids in V583. It contains 19 genes, of which 16 were represented on our microarrays. The highest content of pTEF3 genes was found in strain NCDO 581 (eight genes), while the number of present pTEF3 genes in the other strains varied from zero (V24) to six (179Vet) (Fig. 5). Thus, the environmental strain V24 contained the lowest number of genes similar to the plasmid genes of V583, while multiresistant strain 179Vet and strain NCDO 581 (type strain) harbored a high percentage of genes from all three of the V583 plasmids. The plasmid contents of the nine strains examined here should be examined further.

Conclusions. At present, only one complete genome sequence of the species *E. faecalis* is available. Thus, we do not know yet how representative the V583 genome sequence is of the species as a whole. Considering the fact that V583 contains a huge amount of MGEs and exogenously acquired DNA, it is probable that the V583 genome is larger than the genomes of, e.g., environmental *E. faecalis* strains. This hypothesis can be tested e.g., by an approximate genome size determination of the test strains by pulsed-field gel electrophoresis or, alternatively, by using the ArrayOME approach recently described by Ou et al. (22). What also limits the results of studies with CGH using single-genome arrays is the fact that they only enable us to do one-way comparisons of the strains; genes that are present in the test strains only will not be discovered in these studies. For further insight into the exact gene contents of the test strains, primer walking and sequencing starting from interesting present genes may be done. The inherent noise in microarray experiments and data should also be mentioned. Microarrays seldom provide data as exact as those provided by many traditional single-gene techniques. But microarrays and CGH present us with an overwhelming amount of data that show important evolutionary trends. The information obtained by CGH is crucial, providing us with extensive gene content information that cannot be obtained with traditional techniques. Detailed information on single genes can be obtained, e.g., by simple PCRs.

Regarding the species *E. faecalis*, our CGH analysis has revealed great genome plasticity and diversity. Most of this variation is related to the V583 mobilome, and with regard to, e.g., sugar metabolism, the variation among strains is much lower. The CGH analysis revealed various contents of exogenously acquired DNA (antibiotic resistance- and virulence-related genes) in the test strains, which support the theory of independent selection of antibiotic resistance and virulence traits (10) during the evolution of the species. To define the core of functional genes defining the species *E. faecalis*, more CGH studies are needed. Our studies, along with the work by Lepage et al. (18), have identified a group of genes that probably is a part of the *E. faecalis* core genes. It is hoped that future CGH studies will also reveal the reason why some *E. faecalis* strains can be used as probiotics or in starter cultures while others are feared opportunistic pathogens.

ACKNOWLEDGMENTS

This work was financially supported by grants from the Research Council of Norway to I.F.N.

We thank Zhian Salehian for assistance with the API 50 CH assay. We thank Henning Sørum, The Norwegian School of Veterinary Science, for kindly providing the 138Vet and 179Vet strains.

REFERENCES

1. **Aakra, A., H. Vebo, L. Snipen, H. Hirt, A. Aastveit, V. Kapur, G. Dunny, B. E. Murray, and I. F. Nes.** 2005. Transcriptional response of *Enterococcus faecalis* V583 to erythromycin. Antimicrob. Agents Chemother. **49:**2246–2259.
2. **Behr, M. A., M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane, and P. M. Small.** 1999. Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science **284:**1520–1523.
3. **Bensing, B. A., and G. M. Dunny.** 1993. Cloning and molecular analysis of genes affecting expression of binding substance, the recipient-encoded receptor(s) mediating mating aggregate formation in *Enterococcus faecalis*. J. Bacteriol. **175:**7421–7429.
4. **Cummings, C. A., M. M. Brinig, P. W. Lepp, S. van de Pas, and D. A. Relman.** 2004. *Bordetella* species are distinguished by patterns of substantial gene loss and host adaptation. J. Bacteriol. **186:**1484–1492.
5. **Domig, K. J., H. K. Mayer, and W. Kneifel.** 2003. Methods used for the isolation, enumeration, characterisation and identification of *Enterococcus* spp. 1. Media for isolation and enumeration. Int. J. Food Microbiol. **88:**147–164.
6. **Domig, K. J., H. K. Mayer, and W. Kneifel.** 2003. Methods used for the isolation, enumeration, characterisation and identification of *Enterococcus* spp. 2. Pheno- and genotypic criteria. Int. J. Food Microbiol. **88:**165–188.
7. **Dziejman, M., E. Balon, D. Boyd, C. M. Fraser, J. F. Heidelberg, and J. J. Mekalanos.** 2002. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc. Natl. Acad. Sci. USA **99:**1556–1561.
8. **Frost, L. S., R. Leplae, A. O. Summers, and A. Toussaint.** 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. **3:**722–732.
9. **Fukiya, S., H. Mizoguchi, T. Tobe, and H. Mori.** 2004. Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. J. Bacteriol. **186:**3911–3921.
10. **Gilmore, M. S., and J. J. Ferretti.** 2003. The thin line between gut commensal and pathogen. Science **299:**1999–2002.
11. **Hancock, L., and M. Perego.** 2002. Two-component signal transduction in *Enterococcus faecalis*. J. Bacteriol. **184:**5819–5825.
12. **Hancock, L. E., and M. Perego.** 2004. Systematic inactivation and phenotypic characterization of two-component signal transduction systems of *Enterococcus faecalis* V583. J. Bacteriol. **186:**7951–7958.
13. **Hinchliffe, S. J., K. E. Isherwood, R. A. Stabler, M. B. Prentice, A. Rakin, R. A. Nichols, P. C. Oyston, J. Hinds, R. W. Titball, and B. W. Wren.** 2003. Application of DNA microarrays to study the evolutionary genomics of *Yersinia pestis* and *Yersinia pseudotuberculosis*. Genome Res. **13:**2018–2029.
14. **Kilpper-Bälz, R., and K. H. Schleifer.** 1984. Nucleic acid hybridization and cell wall composition of pyogenic streptococci. FEMS Microbiol. Lett. **24:**355–364.
15. **Klein, G.** 2003. Taxonomy, ecology and antibiotic resistance of enterococci from food and the gastro-intestinal tract. Int. J. Food Microbiol. **88:**123–131.
16. **Koide, T., P. A. Zaini, L. M. Moreira, R. Z. Vencio, A. Y. Matsukuma, A. M. Durham, D. C. Teixeira, H. El-Dorry, P. B. Monteiro, A. C. da Silva, S. Verjovski-Almeida, A. M. da Silva, and S. L. Gomes.** 2004. DNA microarray-based genome comparison of a pathogenic and a nonpathogenic strain of *Xylella fastidiosa* delineates genes important for bacterial virulence. J. Bacteriol. **186:**5442–5449.
17. **Lauková, A., S. Czikkova, Z. Vasilkova, P. Juris, and M. Marekova.** 1998. Occurrence of bacteriocin production among environmental enterococci. Lett. Appl. Microbiol. **27:**178–182.
18. **Lepage, E., S. Brinster, C. Caron, C. Ducroix-Crepy, L. Rigottier-Gois, G. Dunny, C. Hennequet-Antier, and P. Serror.** 2006. Comparative genomic hybridization analysis of *Enterococcus faecalis*: identification of genes absent from food strains. J. Bacteriol. **188:**6858–6868.
19. **Lindsay, J. A., C. E. Moore, N. P. Day, S. J. Peacock, A. A. Witney, R. A. Stabler, S. E. Husain, P. D. Butcher, and J. Hinds.** 2006. Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. J. Bacteriol. **188:**669–676.
20. **Novak, R., B. Henriques, E. Charpentier, S. Normark, and E. Tuomanen.**

1999. Emergence of vancomycin tolerance in *Streptococcus pneumoniae*. Nature **399:**590–593.

21. **Oliver, D. R., B. L. Brown, and D. B. Clewell.** 1977. Analysis of plasmid deoxyribonucleic acid in a cariogenic strain of *Streptococcus faecalis*: an approach to identifying genetic determinants on cryptic plasmids. J. Bacteriol. **130:**759–765.
22. **Ou, H. Y., R. Smith, S. Lucchini, J. Hinton, R. R. Chaudhuri, M. Pallen, M. R. Barer, and K. Rajakumar.** 2005. ArrayOme: a program for estimating the sizes of microarray-visualized bacterial genomes. Nucleic Acids Res. **33:**e3.
23. **Paulsen, I. T., L. Banerjei, G. S. Myers, K. E. Nelson, R. Seshadri, T. D. Read, D. E. Fouts, J. A. Eisen, S. R. Gill, J. F. Heidelberg, H. Tettelin, R. J. Dodson, L. Umayam, L. Brinkac, M. Beanan, S. Daugherty, R. T. DeBoy, S. Durkin, J. Kolonay, R. Madupu, W. Nelson, J. Vamathevan, B. Tran, J. Upton, T. Hansen, J. Shetty, H. Khouri, T. Utterback, D. Radune, K. A. Ketchum, B. A. Dougherty, and C. M. Fraser.** 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. Science **299:**2071–2074.
24. **Poly, F., D. Threadgill, and A. Stintzi.** 2004. Identification of *Campylobacter jejuni* ATCC 43431-specific genes by whole microbial genome comparisons. J. Bacteriol. **186:**4781–4795.
25. **Quackenbush, J.** 2002. Microarray data normalization and transformation. Nat. Genet. **32**(Suppl.)**:**496–501.
26. **Robertson, G. T., J. Zhao, B. V. Desai, W. H. Coleman, T. I. Nicas, R. Gilmour, L. Grinius, D. A. Morrison, and M. E. Winkler.** 2002. Vancomycin tolerance induced by erythromycin but not by loss of *vncRS*, *vex3*, or *pep27* function in *Streptococcus pneumoniae*. J. Bacteriol. **184:**6987–7000.
27. **Sahm, D. F., J. Kissinger, M. S. Gilmore, P. R. Murray, R. Mulder, J. Solliday, and B. Clarke.** 1989. In vitro susceptibility studies of vancomycin-resistant *Enterococcus faecalis*. Antimicrob. Agents Chemother. **33:**1588–1591.
28. **Salama, N., K. Guillemin, T. K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow.** 2000. A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. Proc. Natl. Acad. Sci. USA **97:**14668–14673.
29. **Shankar, N., A. S. Baghdayan, and M. S. Gilmore.** 2002. Modulation of virulence within a pathogenicity island in vancomycin-resistant *Enterococcus faecalis*. Nature **417:**746–750.
30. **Snipen, L., D. Repsilber, L. Nyquist, A. Aakra, A. Ziegler, and A. Aastveit.** 2006. Detection of divergent genes in microbial aCGH experiments. BMC Bioinformatics **7:**181.
31. **Stiles, M. E., and W. H. Holzapfel.** 1997. Lactic acid bacteria of foods and their current taxonomy. Int. J. Food Microbiol. **36:**1–29.
32. **Taboada, E. N., R. R. Acedillo, C. D. Carrillo, W. A. Findlay, D. T. Medeiros, O. L. Mykytczuk, M. J. Roberts, C. A. Valencia, J. M. Farber, and J. H. Nash.** 2004. Large-scale comparative genomics meta-analysis of *Campylobacter jejuni* isolates reveals low level of genome plasticity. J. Clin. Microbiol. **42:**4566–4576.
33. **Tomich, P. K., F. Y. An, and D. B. Clewell.** 1980. Properties of erythromycin-inducible transposon Tn*917* in *Streptococcus faecalis*. J. Bacteriol. **141:**1366–1374.
34. **Trotter, K. M., and G. M. Dunny.** 1990. Mutants of *Enterococcus faecalis* deficient as recipients in mating with donors carrying pheromone-inducible plasmids. Plasmid **24:**57–67.